Articles

Thirteen- and 18-month-old Infants Recognize When They Need Referential Information

Amrisha Vaish, *Max Planck Institute for Evolutionary Anthropology*, Özlem Ece Demir, *University of Chicago* and Dare Baldwin, *University of Oregon*

Abstract

To learn from conspecifics, infants would be greatly advantaged by knowing when to seek information from them. Although in prior work infants used a labeler's gaze direction to infer the referent of a novel label, it was unclear whether infants in these studies recognized that they needed information or were happening upon the information by simply orienting to the labeler's voice. To address this issue, we presented 13- and 18-month-olds with either one or two novel objects and provided a novel label. If infants seek referential information, they should look more to the labeler in the presence of two objects relative to one object, because the labeler's intended referent is highly ambiguous in the two-object case. This prediction was confirmed in two studies. In contrast, infants' looking was equivalent in the presence of one vs. two objects in a baseline phase, when no labels were provided. Thus, 13- and 18-month-olds actively seek clarifying gaze information to resolve ambiguous learning situations. Word learning appears to be a rich pragmatic process as early as the end of the first year of life.

Keywords: information seeking; referential understanding; gaze following; word learning

Introduction

The ability to gather various types of information from conspecifics is thought to play a central role in our acquisition of world knowledge and in enabling the 'cultural learning' that is thought to be responsible for uniquely human achievements such as language and science (Baldwin & Moses, 1996; Csibra & Gergely, 2006; Tomasello, 1999). Such information gathering is especially important for young infants, who must and do learn entire systems of knowledge such as one or more languages within just a few years of life. However, in order to successfully learn these systems, infants must be able to recognize the circumstances under which they need to consult others for information (Baldwin & Moses, 1996). In this article, we focus on this ability in

infants. Specifically, by examining when infants reference informers, we attempt to better understand how skilled infants are at gathering social information.

One especially informative referential cue is gaze direction: If infants are to learn the correct referents of new information, such as novel words, then following the informer's gaze to the appropriate referent is crucial for making the correct associations (Baldwin, 1993). If infants did not actively use gaze direction, they would likely often make mistakes in word–world mappings. In two studies addressing this issue, Baldwin (1991, 1993) assessed what types of information infants use to guide their inferences about word meanings. In these studies, while infants were exploring one novel object, the experimenter looked at and provided a novel label for a different novel object. Baldwin (1991, 1993) found that infants actively consulted the speaker's gaze to help them negotiate the complexities of this 'discrepant labeling' situation. Specifically, on hearing the novel label, 18–19-month-olds (and, to a lesser degree, 16–17-month-olds), checked the speaker's face and noted her gaze direction and, when later probed for comprehension of the novel label, they systematically selected the object on which the experimenter had been focused while labeling rather than the object on which they themselves had been focused. Moses, Baldwin, Rosicky, and Tidball (2001) reported similar findings with 12- and 18-month-olds in a social referencing context.

These studies show that infants as young as 12 to 18 months attend to the referential intent of informers. However, it remains unclear whether infants in these studies were truly looking for information from the experimenter. One plausible alternative is that infants were looking to the experimenter as an orienting response to her voice and then receiving crucial referential information by happenstance when they looked (Baldwin & Moses, 1996). If this explanation is correct, then infants may not be expert social information gatherers and might still have to develop the ability to identify ambiguous situations in which they need clarifying information. On the other hand, infants in these prior studies may well have been actively seeking disambiguating information, pointing to a striking degree of skill in recognizing situations where socio-pragmatic information is key to learning success, and actively capitalizing on such information. The present two studies were designed to test between these two possibilities.

Although these two studies were conducted independently by two different groups of researchers in different labs and thus varied somewhat in their procedures, the general procedure of both studies was as follows: an experimenter presented 13- and 18-month-old infants with either one or two novel objects, and then produced a novel label. The presence of two novel objects should engender referential ambiguity (as in Baldwin, 1991, 1993), whereas when only one novel object is present, there should be little referential ambiguity. Thus, if infants truly seek information about the novel label's referent, they should look more to the labeler in the presence of two novel objects than in the presence of one. We therefore predicted that 18-month-olds, and perhaps even 13-month-olds, would look to the labeler more in the presence of two novel objects, but only in response to a novel label (not during a non-labeling baseline). Both studies also included comprehension trials to assess infants' ability to infer the novel label's referent.

Study 1

Method

Participants. Thirty-six full-term infants from the Chicago area participated in this study. Infants' parents had been contacted by advertisements or mailings, and were

offered $10 as travel reimbursement. An inclusion criterion was that infants heard English at least 50 percent of the time (which parents confirmed on the telephone when making an appointment). Eighteen infants were 13-month-olds (9 girls; $M = 12$ months, 22 days; $SD = 21.9$ days), and 18 were 18-month-olds (10 girls; $M = 17$ months, 24 days; $SD = 7.1$ days). Participants were ethnically diverse: 44.44 percent White, 16.67 percent Hispanic, 27.78 percent African-American, 2.78 percent Asian, and 8.33 percent of mixed ethnicity. An additional 12 infants were tested but were excluded due to experimenter error (6), fussiness (2), hearing English less than 50 percent of the time (2), equipment failure (1), and parent not following instructions (1).

Materials

Stimuli. The same two novel toys were used in both the Baseline and Labeling phases (see description of the phases below). One novel toy was a small red plastic tube, open at one end and with bumps along one long edge. The second was a coaster with a purple rim and a brown, cork surface. A prior preference test showed that infants had no systematic preference for either toy. Parents also picked from among six familiar toys (shoe, block, teddy bear, bunny, frog, and duck), two toys for which their infant would know the labels.

Novel label. The novel label used (*toma*) was selected to be highly pronounceable and consistent with English phonotactic constraints (Baldwin, 1991).

Equipment. Two video cameras captured close-up views of infants and experimenter 1 (E1). Both images were recorded on a video recorder with frame-by-frame viewing capability.

Design. We used a $2 \times 2 \times 3$ mixed-design with two between-subjects variables (condition: one vs. two object(s); age: 13- and 18-month-olds) and one within-subjects variable (trial type: baseline, labeling 1, and labeling 2). Infants in the one object condition were shown one out-of-reach novel object and infants in the two objects condition were shown two out-of-reach objects. The baseline was always conducted first, followed by two labeling trials. The two phases differed only with respect to whether the experimenter produced novel labels during the trials or not. Following labeling trials, all infants received two comprehension trials.

Procedure. After warming up with two experimenters (E1 and experimenter 2 [E2]), infants were seated in a booster seat at a table. Caregivers sat at a different table out of infants' view and, throughout the procedure, filled out the short version of the Mac-Arthur communicative development inventories (MCDI) level I (Fenson et al., 2000). E1 sat across the table from infants, at a distance of .8 m. E2 stood behind infants and kept time.

Baseline trials (20 s). E1 presented infants with a small tray holding either one or both novel objects. In the one object condition, one object was placed at one end of the tray (object and side counterbalanced across infants), and the second object was out of infants' sight; in the two objects condition, the two objects were placed at either ends of the tray (sides counterbalanced across infants). Objects were thus never placed at the center of the tray (to match labeling trials; the reason for doing this in labeling trials

will be explained shortly). After placing the tray in front of infants but out of their reach, E1 silently looked at the center of the tray without looking at an object for 20 s.

After baseline, E1 gave both novel objects to infants so that, regardless of condition, infants would be familiar with both objects. For 20 s, E1 ensured that infants interacted approximately equally with both objects, after which she put the novel objects out of view. Next, to familiarize infants with the procedure of the comprehension trials to be conducted later, E1 asked infants to put one of two familiar toys (that caregivers had indicated infants knew), in a bucket. If infants chose correctly, E1 clapped and cheered. If infants chose incorrectly, E1 corrected them and repeated her request. Only once infants had chosen correctly did E1 put away the familiar toys and start the next phase.

Labeling trials (2 trials, 20 s each). In the first labeling trial, E1 placed either one or both novel objects (depending on condition) at the end(s) of the tray exactly as in baseline, and placed the tray at the same place as during baseline. After ensuring that infants were paying attention to her, she produced a novel label three times ('Look, a toma! See? It's a toma! Look, that's a toma'), while alternating gaze between infants and the center of the tray. She then gazed silently at the center of the tray for 20 s. The second labeling trial was identical except that the object locations were switched.

Importantly, because in neither condition was there an object at the center of the tray, infants could not instantly connect E1's gaze with an object, which meant they had to look to her more to decipher the object of her gaze. In this case, if infants are seeking information, they should look to E1 more when there are two objects present because E1's gaze direction is the critical cue to deciphering which object she is labeling; when only one object is present, infants should look less to her as they can infer that E1 *must* be labeling the only object present even if she is not gazing directly at it.

Comprehension trials. Finally, E1 presented infants with both novel objects and asked them to put the toma in the bucket. If infants did not respond, E1 asked them to point to or give her the toma, and if that was ineffective, then to give the toma to the caregiver. The trial ended when infants selected an object or after 1 min had expired. The second comprehension trial was identical except that the object locations were switched.

Coding and Reliability. In baseline and labeling trials, 20 s of silence seemed too long as many infants were restless in the last 5–10 s. We thus used only the first 10 s of each trial. The primary coder used Interact (Version 8.0, 2007) to code duration and number of infants' looks to E1 and the object(s) during baseline and labeling. A reliability coder (blind to condition and hypotheses) coded a randomly selected eight infants (four at each age). Inter-rater agreement was excellent: $\kappa = .83$ and $\kappa = .85$ for 18- and 13-month-olds, respectively.

The Comprehension trials were coded only for infants in the one object condition (nine at each age), as only in this condition could infants make a correct response, that is, choose the one object that had been present during labeling. Infants' choices were coded as follows: first object pointed to; if infants did not point to an object, then first object touched; if infants touched both objects simultaneously, then first object given to E1 or the caregiver or put in the bucket. The reliability coder coded a randomly selected four infants (two at each age). Inter-rater agreement regarding which object children chose was 100 percent.

Results

We first present infants' MCDI scores, followed by analyses of the duration and number of infants' looks to E1. We then present analyses of the duration and number of infants' looks to the object(s), which were conducted to assess infants' attention to one vs. two objects and to help interpret the findings of infants' looks to E1. (For Studies 1 and 2, looks to the experimenter and the object(s) are summarized in Table 1). Finally, we present analyses of infants' comprehension performance.

MCDI. As expected, 18-month-olds comprehended ($M = 57.88$; $SD = 17.99$, *range*: 20–89) and produced ($M = 29.00$; $SD = 16.00$; *range*: 6–53) significantly more words than 13-month-olds (comprehension: $M = 28.06$; $SD = 16.98$; *range*: 5–82; production: $M = 7.89$; $SD = 10.27$; *range*: 0–43), both $ts (32) > 4.51$, both $ps < .0005$. Comprehension and production scores did not differ across gender or condition, all $ps > .188$. Note also that for the following analyses, including infants' comprehension and production scores as covariates did not change the results; we thus only present results of analyses without these variables as covariates.

Looking to the Experimenter. Our central question concerned whether infants looked more to the labeler when reference was ambiguous (two objects condition) vs. relatively unambiguous (one object condition). If meaningful, this difference should emerge during the labeling but not the baseline trials. The primary question was thus whether a condition (one vs. two object(s)) × trial type (baseline vs. labeling) interaction would emerge. Gender was not included in these analyses as preliminary analyses revealed no effects of gender.

Duration of looks. A mixed-design analysis of variance (ANOVA) using total duration of infants' looks to E1 as a dependent measure was conducted with trial type (three levels: baseline, labeling 1, labeling 2) as a within-subjects factor and condition (one object vs. two objects) and age group (13- vs. 18-month-olds) as between-subjects factors. This analysis revealed a main effect of trial type, $F(2, 64) = 4.01$, $p = .023$, pη^2 = .11: infants looked longer to E1 in Baseline trials ($M = 3.30$ s, $SD = 2.18$) than in labeling 1 ($M = 3.08$ s, $SD = 2.43$), and longer in labeling 1 than labeling 2 ($M = 2.92$ s, $SD = 2.12$). However, pairwise comparisons (with Bonferroni correction, $\alpha = .017$) revealed no differences, all $ps > .036$.

As predicted, this main effect of trial type was qualified by a significant interaction between trial type and condition, $F(2, 64) = 5.14$, $p = .009$, p$\eta^2 = .14$ (see Figure 1). This interaction was analyzed further using independent samples *t*-tests (with Bonferroni correction, $\alpha = .017$). As expected, in baseline, no difference emerged in total duration of looks to E1 in two objects ($M = 4.33$ s, $SD = 2.07$) vs. one object conditions ($M = 3.48$ s, $SD = 1.69$, $t(34) = 1.36$, $p = .183$), whereas in labeling 1, infants looked significantly longer to E1 in the two objects ($M = 4.09$ s, $SD = 2.32$) than in the one object condition ($M = 2.08$ s, $SD = 2.14$, $t(34) = 2.711$, $p = .010$). This difference did not emerge in labeling 2, $p = .614$. Infants at both ages thus looked longer to the labeler in the two objects condition (at least in the first labeling trial), presumably to gather more information about the ambiguous situation. There were no other significant effects.

Number of looks. A similar mixed-design ANOVA examined total number of looks to E1. There was no significant interaction, $p = .32$, but a main effect of trial type, $F(1,

Table 1. *Means* (and *SDs*) of the Looking Measures of Study 1 and Study 2

Measure	Study 1				Study 2			
	One object		Two objects		One object		Two objects	
	Baseline	Labeling 1	Baseline	Labeling 1	Baseline	Labeling	Baseline	Labeling
Looking to the experimenter								
Duration of looks(s)	3.48	2.08	4.33	4.09	1.5	3.1	2.5	5.0
	(1.69)	(2.14)	(2.07)	(2.32)	(2.3)	(3.3)	(2.5)	(4.0)
Number of looks	2.67	1.83	2.61	2.11	1.8	2.3	2.1	4.5
	(1.08)	(.99)	(1.04)	(.83)	(1.9)	(1.8)	(1.5)	(2.9)
Looking to the object(s)								
Duration of looks(s)	3.13	1.24	3.96	2.22	20.0	18.7	19.7	19.5
	(2.06)	(1.37)	(2.35)	(1.45)	(7.1)	(5.1)	(6.9)	(5.7)
Number of looks	2.17	1.22	2.39	1.78	3.2	3.5	3.6	5.1
	(1.10)	(1.17)	(1.20)	(.65)	(1.8)	(1.9)	(1.9)	(2.4)

Note: For Study 2, 'Looks to the object(s)' only includes looks to the target object (the object in infants' possession), not looks to the second object.
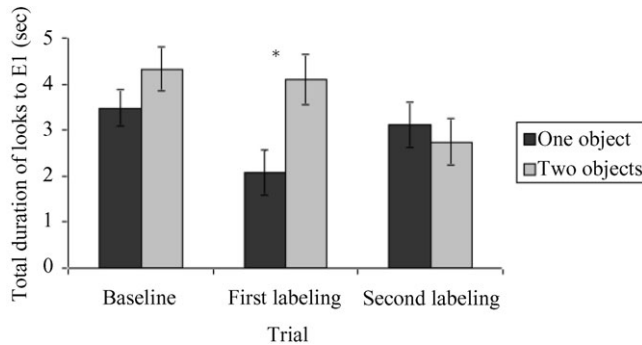
Figure 1. Total Duration of Infants' Looks to the Labeler in Each Condition in Baseline and Labeling Trials in Study 1.
* $p = .01$.

64) $= 9.20$, $p < .0005$, $p\eta^2 = .22$. Pairwise comparisons (with Bonferroni correction, $\alpha = .017$) revealed significantly more looks in the baseline ($M = 2.64$, $SD = 1.05$) than in labeling 1 ($M = 1.97$, $SD = .91$), and more looks in the baseline than in labeling 2 ($M = 1.83$, $SD = 1.03$), both ps $= .003$. These effects, and the similar (but non-significant) effects seen in the durations of infants' looks, might be due to fatigue, given that labeling trials always followed baseline. Number of looks did not differ across labeling 1 and 2, $p = .387$. There were no other significant effects.

Looking to the Objects.

Duration of looks. A mixed-design ANOVA was conducted with total duration of infants' looks to the object(s) as the dependent measure, trial as a within-subjects factor, and condition and age group as between-subjects factors. This revealed a main effect of trial, $F(2, 64) = 20.22$, $p < .0005$, $p\eta^2 = .387$: infants looked significantly longer at the object(s) during baseline ($M = 3.54$ s, $SD = 2.22$) than during labeling 1 ($M = 1.73$ s, $SD = 1.48$, $p < .0005$) and labeling 2 ($M = 1.34$ s, $SD = 1.60$, $p < .0005$), whereas labeling 1 and labeling 2 did not differ, $p = .230$. There was also a main effect of condition, $F(1, 32) = 5.05$, $p = .032$, $p\eta^2 = .136$: infants looked more at the object(s) in the two objects ($M = 2.66$ s, $SD = 2.12$) than in the one object condition ($M = 1.75$ s, $SD = 1.83$), which would be expected given there were more objects to look at in the two objects case. There were no further significant effects.

Number of looks. A similar mixed-design ANOVA examined total number of infants' looks to the object(s). This revealed only a main effect of trial, $F(2, 64) = 20.63$, $p < .0005$, $p\eta^2 = .392$: infants looked significantly more often at the object(s) during baseline ($M = 2.28$, $SD = 1.14$) than labeling 1 ($M = 1.50$, $SD = .97$, $p < .0005$) and labeling 2 ($M = 1.06$ s, $SD = .83$, $p < .0005$), and more often during labeling 1 than labeling 2, $p = .018$. There were no further significant effects.

Comprehension Trials. The comprehension trials helped assess whether infants made appropriate inferences about the referent of the novel label. That is, given the condition differences in duration of infants' looks to the labeler, it seems that infants in the one

object condition needed less information about the novel label's referent. In this case, these infants may have linked the novel label with the only object present despite looking less to the labeler. To assess comprehension, proportion correct across the two comprehension trials in the one object condition was calculated, and the mean proportions compared against chance (.50). As a group, infants did not score above chance, $p = .542$. However, 18-month-olds scored nearly above chance ($M = .72$, $SD = .36$, one-sample $t(8) = 1.84$, one-tailed $p = .052$), whereas 13-month-olds did not ($M = .39$, $SD = .33$, $t(8) = -1.00$, one-tailed $p = .174$). Thus, the 18-month-olds in the one object condition, despite looking less to the labeler than infants in the other condition, showed a tendency to link the novel label to the one object present. This supports the idea that infants looked to the labeler to gather information, and that infants in the one object condition looked less because the labeling situation they faced was less ambiguous.

Discussion

In Study 1, uttering a novel label in a referentially ambiguous context—two novel objects—elicited longer looking to the speaker than doing so in a relatively unambiguous context—one novel object. Importantly, no significant differences emerged in the presence of one vs. two novel objects in a baseline phase when no novel labels were uttered. These results suggest that infants recognize when they require disambiguating information, and they seek out this information by referencing the informer.

Note that from baseline to labeling 1, the duration of looks to the experimenter decreased in the one object condition rather than increasing in the two objects condition. This is not surprising given that once infants had seen the novel object(s) in the baseline, their subsequent interest in the objects and in the situation in general would be expected to decline due to boredom or familiarization. Indeed, infants' looks to the object(s) did decline from baseline to labeling trials. In the one object condition, infants' looks to the speaker also followed this expected pattern. In the two objects condition, however, hearing a novel label caused infants to increase their looking to the speaker relative to the one object condition, potentially in order to work out which object she was labeling. Thus, it is plausible that from baseline to labeling 1, infants' looking to the speaker would typically have declined (as it did in the one object case), and that in the two objects case, the need for information raised the duration of looking to the baseline level. Counterbalancing the order of the baseline and labeling trials, as done in Study 2, may address this issue.

The comprehension trials indicated that 18-month-olds in the one object condition showed a tendency to infer a word–object link. This inference is quite sophisticated given that infants looked *less* to the labeler in the one object condition, and the labeler did not look at the object while labeling. Interestingly, and contrary to some prior work (Woodward, Markman, & Fitzsimmons, 1994), 13-month-olds did not make this inference. This may have been because the labeling situation was quite complex because, in the absence of referent-directed gaze, infants had to infer that the label *must* refer to a certain referent given the pragmatics of the situation. Also, because the sample sizes in Study 1 were rather small (per condition, there were only nine infants of each age), the age differences in comprehension performance and the results in general should be considered with some caution.

Study 1 was a first step toward establishing that 13–18-month-olds recognize when they need referential information. Study 2 addressed the same questions as Study 1.

However, Study 2 included a significantly larger sample than Study 1, which proved helpful in addressing some of the limitations of Study 1. Moreover, as Study 2 was conducted independently in a different lab, it served as a test of the robustness of the results of Study 1. Study 2 also used a somewhat modified procedure; the details of the modifications are highlighted below.

Study 2

Method

Participants. The final sample consisted of 64 infants living in a Pacific Northwest university town or its immediate environs, 32 in each age group [13 months ($M =$ 12 months, 27 days, $SD = 14.6$ days) and 18 months ($M = 18$ months, 27 days; $SD = 15.9$ days)]. Half the infants in each age group were girls. Only infants hearing at least 95 percent English at home (which parents confirmed on the telephone when making an appointment), participated in the research. Data from an additional 40 infants were eliminated due to fussiness (16), the infant happening to look away from the target object just as the experimenter began producing a novel label (14), experimenter error (5), parental over-involvement (4), and equipment failure (1). Participants were drawn from a relatively homogeneous population consisting primarily of White, middle-class families, and were recruited via birth announcements published in local papers.

Materials.

   Stimuli. Two pairs of novel toys (a retractable tape measure and a disc encircled by suction cups; a collapsible cup and a toy with clacking bobbles) were utilized in the baseline and labeling phases (see description of the phases below). The novel toys were selected to be lacking a known label to infants between 12 and 19 months, manipulable, visually distinct from one another, equally salient, and sturdy in the face of constant cleansing. A prior preference test showed that infants had no systematic preferences regarding the novel toys within a given pair. Three pairs of familiar toys were used in the comprehension phase (see description below). These were selected based on words that parents indicated (during the scheduling phone call), their infants comprehended, and were balanced in salience within a pair, visually distinctive, manipulable, and sturdy.

   Novel labels. The novel labels (*modi* and *dawnoo*) were selected to be pronounce-able, distinct from one another, and consistent with English phonotactic constraints.

   Equipment. Two video cameras (one with a stopwatch function and the other with frame-by-frame viewing capability), recorded infants' and the experimenter's behavior during the session.

Design. Study 2 used a $2 \times 2 \times 2$ mixed-design, with condition (one vs. two object(s)) and age (13- and 18-month-olds) as between-subjects variables and trial type (baseline vs. labeling) as a within-subjects variable. Although Study 2 was overall quite similar to Study 1, three procedural changes made Study 2 more similar to Baldwin's (1991, 1993) studies: (1) rather than the novel object(s) being placed out of reach during the trials, infants were given one novel object ('target') to examine in both conditions, while a second novel object ('distractor') was placed

within infants' view but out of their reach in the two objects condition; (2) during labeling trials, the experimenter ensured that infants were looking at the target object at the time that she produced the labels; and (3) familiar label comprehension trials were included. Familiar label comprehension performance can provide valuable information, especially if infants perform poorly on novel label comprehension trials, because it helps to clarify whether infants did not register the new word or whether they had difficulty with the comprehension task itself. Four comprehension questions were asked for each label (as opposed to the two used in Study 1), to assess the systematicity of infants' comprehension of a given label. Further procedural modifications are detailed below.

Procedure. After warming up in the playroom, infants were seated in a high chair at a table while parents sat to the side and, throughout the session, filled out the full version of the MCDI questionnaire (rather than the short version used in Study 1). Parents were instructed to refrain from initiating interaction with infants, labeling toys, or displaying any reactions to the events that occurred. Infants were given a short familiarization with the comprehension test format using two familiar toys (usually a dog and a ball) that had been available during warm-up: the experimenter asked infants to find one or the other toy and cheered enthusiastically when infants touched or picked up the correct toy.

Baseline trials. A schematic of the Study 2 procedure is presented in Figure 2. During Baseline trials, the experimenter displayed either one or both novel toys (depending on condition) and presented the target object to the infant. She then sat silently looking at her 'work'—a clipboard with paper that she held just visible beyond the table across from infants. Thus, the experimenter did not look at any novel object during baseline trials, unless infants dropped a toy and the experimenter needed to retrieve it for them.

Labeling trials. The procedure was identical to that of baseline trials except that the experimenter produced a novel label (e.g., 'A modi!') three times while looking at her work. Prior to each labeling utterance, the experimenter waited until the infant was focused on the target object.

Note that only one labeling trial was included, in contrast to the two trials in Study 1. Also, unlike Study 1, the order of trial type (whether baseline or labeling occurred first) was counterbalanced across infants. The aim was for the baseline and labeling trials to last 20 s each. However, because the methodological design required infants to be looking at the target toy when the label was produced, and as it was sometimes difficult to have infants re-focus on the toy, trials occasionally lasted longer than 20 s. Still, when possible, the lengths of the baseline and labeling trials were matched. That is, if the labeling trial occurred first and exceeded 20 s, then the baseline trial's length was matched to the length of the preceding labeling trial. A *post hoc* paired *t*-test revealed that baseline ($M = 26.0$ s, $SD = 6.0$) and labeling ($M = 27.2$ s, $SD = 8.1$) trials did not differ significantly in length, $p > .30$.

In both conditions, at the end of the trial, infants were handed the second novel object (which had thus far been out of view in the one object condition or out of reach in the two objects condition) in order to familiarize them with it prior to the comprehension trials. No labels were uttered during this final familiarization period. The assignment of novel object pairs, which novel object in a pair was the target object, and which novel label was used, was counterbalanced across infants.
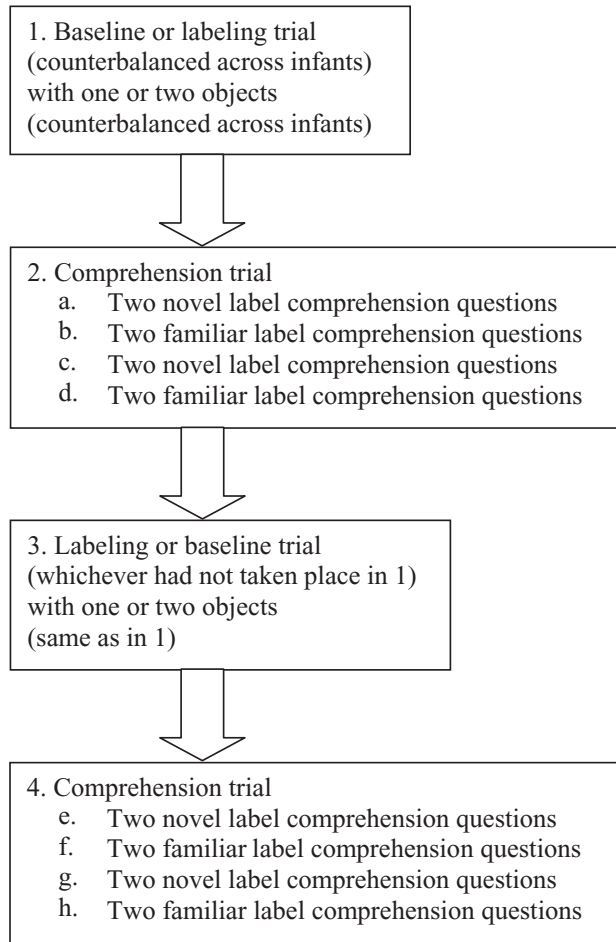
```
┌─────────────────────────────────────┐
│ 1. Baseline or labeling trial        │
│ (counterbalanced across infants)     │
│ with one or two objects              │
│ (counterbalanced across infants)     │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────────────────────┐
│ 2. Comprehension trial                               │
│    a.  Two novel label comprehension questions       │
│    b.  Two familiar label comprehension questions    │
│    c.  Two novel label comprehension questions       │
│    d.  Two familiar label comprehension questions    │
└─────────────────────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│ 3. Labeling or baseline trial        │
│ (whichever had not taken place in 1) │
│ with one or two objects              │
│ (same as in 1)                       │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────────────────────┐
│ 4. Comprehension trial                               │
│    e.  Two novel label comprehension questions       │
│    f.  Two familiar label comprehension questions    │
│    g.  Two novel label comprehension questions       │
│    h.  Two familiar label comprehension questions    │
└─────────────────────────────────────────────────────┘
```

Figure 2. Flow Chart of the Study 2 Procedure.

Comprehension trials. Unlike in Study 1, a comprehension trial followed both the baseline and the labeling trials (thus, all infants participated in two comprehension trials). This helped assess whether the one vs. two object(s) manipulation might have increased infants' selection of the target object simply due to demand characteristics and not due to genuine word comprehension. That is, if greater selection of the target toy occurred in the one object relative to the two objects condition due to genuine label comprehension, this should emerge only after labeling trials (not after baseline trials in which no label had been uttered).

Infants were asked 16 comprehension questions in all: eight in each comprehension trial (four regarding the novel label and four regarding a familiar label). Novel label comprehension questions involved repeatedly presenting the same two novel objects (the target and distractor objects) and asking infants to identify the labeled object. Familiar label comprehension questions involved repeatedly presenting the same two objects with familiar labels (based on parental report) and asking infants to identify the labeled object. During the comprehension trials for a given infant, the labeled object's

location (left vs. right) was counterbalanced. Assignment of familiar toys was also roughly counterbalanced with respect to condition, trial type, and order of trial type, although precise counterbalancing could not be achieved because infants differed in terms of which labels their parents reported them to comprehend. Pairs of novel label comprehension questions alternated with pairs of familiar label questions, always beginning with the former. Various props (a tray, a basket, a cosmetics bag, and a tea cozy), were used to enhance infants' interest in the comprehension test: on each comprehension trial, the experimenter presented the target and distractor objects side by side either on or inside one of these props and asked the comprehension question. Labels were embedded in varied carrier contexts ('Point to the *modi*!', 'Where is the *modi*?', 'Show me the *modi*!' and 'Can you find the *modi*?'). Props and carrier phrases occurred in a prearranged order held constant for all infants.

Coding and Reliability. During baseline and labeling phases, coders recorded the start and stop times of infants' looks to the target object, the experimenter, the distractor object (in the two objects condition), and other (anything else). Independent coding of four infants by two different coders revealed 85 percent agreement regarding location of infants' looks across baseline and labeling trials. Coders' independent judgments of the duration and number of these infants' looks were strongly correlated, $r(49) = .99$, $p = .0001$, and $r(7) = .95$, $p = .0001$, for duration and number, respectively. Comprehension trials were coded for which object (target vs. distractor) infants (a) looked at first, (b) touched first, (c) grasped first, and (d) pointed to first. Infants' choices were coded as follows: object pointed to first; object grasped first if no point was made; object touched first if no grasp was made; if infants touched both objects simultaneously, then object given or shown to the experimenter or placed in the container.

Results

Again, we first present infants' MCDI scores, followed by analyses of the duration and number of infants' looks to the experimenter. We then present analyses of the duration and number of infants' looks to the object(s), and finally, analyses of infants' comprehension performance.

MCDI. As in Study 1, and as expected, infants in the older group comprehended ($M = 236.1$; $SD = 85.3$, *range*: 96–381) and produced ($M = 93.9$; $SD = 76.8$; *range*: 2–281) significantly more words than those in the younger group ($M = 111.2$; $SD = 64.3$; *range*: 29–290 and $M = 12.6$; $SD = 15.2$; *range*: 0–77 for comprehension and production, respectively), ts $(53) > 6.0$, ps $< .001$. Comprehension and production scores did not differ across gender or condition, all ps $> .30$. For the following analyses, using infants' comprehension and production scores as covariates did not change the general pattern of results; we will thus only present results of analyses without these variables as covariates.

Looking to the Experimenter.

   Duration of looks. Preliminary analyses revealed no effects of gender, thus all reported analyses collapsed across gender. A mixed-design ANOVA (with trial type [baseline vs. labeling] as a within-subjects factor and condition [one vs. two objects] and age group [13 vs. 18-month-olds] as between-subjects factors) examined duration
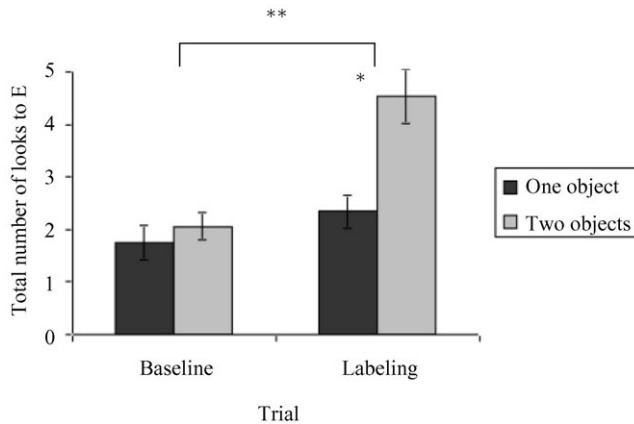
Figure 3. Number of Infants' Looks to the Labeler in Each Condition in Baseline and Labeling Trials in Study 2.
* *p* = .001, ** *p* < .005.

of infants' looks to the experimenter as a dependent variable. Similar to Study 1, infants looked to E for longer when labels were uttered in the presence of two novel objects (*M* = 5.0 s; *SD* = 4.0) relative to one novel object (*M* = 3.1 s; *SD* = 3.3), independent-samples *t*(62) = 2.0, *p* = .05. However, in the ANOVA, only a main effect of trial type was significant, *F*(1, 60) = 16.0, *p* < .0005: infants looked to the experimenter for longer in response to labeling (*M* = 4.1 s; *SD* = 3.7) than during baseline (*M* = 2.0 s; *SD* = 2.5). No other significant effects emerged.

Number of looks. A similar mixed-design ANOVA examined total number of looks to E, revealing a significant main effect of trial type, *F*(1, 60) = 24.05 *p* < .0005, $p\eta^2$ = .29: labeling elicited more looks to the experimenter overall (*M* = 3.44, *SD* = 2.6) than did Baseline (*M* = 1.9, *SD* = 1.7). The analysis also revealed a main effect of condition, *F*(1, 60) = 8.9, *p* = .004, $p\eta^2$ = .13: Overall, infants displayed more looks to the experimenter in the presence of two objects (*M* = 3.3, *SD* = 2.6) than one object (*M* = 2.0, *SD* = 1.8).

Importantly, these main effects were qualified by the predicted interaction between trial type and condition, *F*(1, 60) = 9.05, *p* = .004, $p\eta^2$ = .13 (see Figure 3). Pairwise comparisons (with Bonferroni correction, α = .025) revealed a significantly greater number of looks to the experimenter in the two objects (*M* = 4.5, *SD* = 2.9) than in the one object condition (*M* = 2.3, *SD* = 1.8), independent samples *t*(62) = 3.62, *p* = .001, but no significant difference in the baseline phase in the presence of two (*M* = 2.1, *SD* = 1.5) vs. one novel object (*M* = 1.8, *SD* = 1.9), independent samples *t*(62) = .72, *p* = .47. Paired comparisons (with Bonferroni correction, α = .025) also revealed that labeling promoted looks to the experimenter relative to the baseline phase when two objects were present, *t*(31) = 4.9, *p* < .0005, but not when only one object was present, *t*(31) = 1.6, *p* = .12. Together, these findings clarified that infants intensified their monitoring of the experimenter's face when she uttered a novel label, and especially so when a novel label was uttered in the referentially ambiguous (two objects) case. The ANOVA revealed no other significant effects.

Looking to the Target Object.

Duration of looks. A mixed-design ANOVA was conducted with total duration of infants' looks to the target object as a dependent measure, trial type as a within-subjects factor, and condition and age group as between-subjects factors. This revealed no significant effects.

Number of looks. A similar mixed-design ANOVA examining total number of infants' looks to the target object revealed a main effect of trial, $F(1, 60) = 10.3$, $p = .002$, $p\eta^2 = .15$: Infants looked more often at the target object during labeling ($M = 4.3$, $SD = 2.3$) than during baseline ($M = 3.4$, $SD = 1.9$). A trial type condition interaction also emerged, $F(1, 60) = 4.7$, $p = .033$, $p\eta^2 = .07$: labeling elicited a greater number of looks to the target object in the two objects condition ($M = 5.1$, $SD = 2.4$) relative to the one object condition ($M = 3.5$, $SD = 1.9$), $t(62) = 2.9$, $p = .005$, but differences did not emerge in the baseline trials (one object $M = 3.2$, $SD = 1.8$; two objects $M = 3.6$, $SD = 1.9$; $p = .43$). As well, looks to the target object were more frequent in response to labeling than during baseline in the two objects condition, paired $t(31) = 3.5$, $p = .001$, whereas no such difference emerged in the one object condition. There were no further significant effects.

These findings seem consistent with infants' responding in the comprehension trials described below: when a novel label was provided, infants were inclined to look frequently at the object in their possession, especially when reference was ambiguous (in the two objects condition), because they (a) checked the speaker more frequently under these circumstances, and presumably (b) inferred that the object in their possession was the novel label's referent.

Comprehension Trials. The comprehension trials in Study 2 allowed us to examine several issues. Firstly, because they were conducted after both baseline and labeling phases, they provided basic control information: if the comprehension task indeed assessed learning of the novel label, then infants should perform better on comprehension questions asked after introduction of the label (i.e., after the labeling phase) than those asked without infants having heard the label before (i.e., after baseline). Secondly, one might ask whether infants would display stronger comprehension performance for the novel label when it was introduced in an unambiguous context (one object condition) relative to a referentially ambiguous context (two objects condition).

We tested these issues via a $2 \times 2 \times 2$ mixed-design ANOVA with trial type (baseline vs. labeling) as a within-subjects variable, and condition and age group as between-subjects variables. This analysis revealed a main effect of trial type: infants displayed higher rates of selecting the target object in response to novel label comprehension questions following labeling ($M = .57$; $SD = .33$) than following baseline ($M = .46$; $SD = .31$), $F(1, 60) = 4.6$, $p = .04$. No other significant effects emerged. Most notably, there were (1) no effects of age group, and (2) no significant differences in comprehension performance for novel labels introduced in the one object vs. two objects conditions. Furthermore, infants' selection of the target object exceeded chance levels in response to novel label comprehension questions following labeling, one-sample $t(63) = 1.78$, one-tailed $p = .04$, but not following baseline.

Nicely, separate analyses revealed that infants' comprehension of *familiar* labels did not differ following labeling ($M = .68$; $SD = .26$) vs. baseline ($M = .68$; $SD = .27$). Also, as one would predict, comprehension of familiar labels exceeded chance levels after

both phases, one-sample ts(63) > 5.1, one-tailed ps < .001. The fact that familiar label comprehension was strong after both labeling and baseline phases helps clarify that the differences in novel label comprehension between labeling and baseline were due to genuine label learning in the labeling phase rather than due to confusion or inattentiveness resulting from the baseline phase. That is, having comprehension questions 'sprung' upon them after the silent baseline phase might have confused infants or reduced their attentiveness to the labels in the comprehension questions; however, infants' consistently high comprehension performance with familiar labels following both baseline and labeling phases suggests this was not the case. This finding bolsters the interpretation that the differences in novel label comprehension between the phases were meaningful *vis-à-vis* infants' processing of the meanings of the novel words.

Discussion

Despite methodological differences across studies, the findings from Study 2 corroborated those of Study 1: infants displayed greater looking to the speaker when she produced a novel label in a referentially ambiguous context (two objects condition), than in a less ambiguous context (one object condition). Moreover, this looking difference was specifically elicited by the labeling; during baseline, infants' looking did not differ in the presence of one vs. two novel objects. Strikingly, as in Study 1, 13-month-olds displayed the same pattern of intensified checking of the labeler when reference was ambiguous as older infants did.

   Interestingly, in Study 1, infants' information seeking was manifest in the duration of looks to the speaker, whereas in Study 2, it was especially pronounced in the number of looks. Possible reasons for this difference lie in the procedural details of the studies. In Study 1, after labeling, the speaker gazed at the tray holding the objects, so infants presumably looked to her for longer in the two objects condition in order to work out where, that is, which object, she was looking at (perhaps by trying to follow her gaze). In Study 2, in contrast, the speaker was looking at her 'work' while labeling, so she was unavailable to provide the referential information infants needed; infants might thus have looked to her to check whether she was available rather than where she was looking, and when they needed more information (in the two objects condition), they checked more often whether she was available. Thus, the methodological differences may account for the distinct but related measures that revealed significant differences in the two studies.

   Infants' responses to comprehension questions in Study 2 provided additional information on several fronts. Firstly, infants' better-than-chance comprehension performance with novel labels following the labeling phase indicated that they attended to and processed the novel labels, bolstering the possibility that their looks during labeling were meaningful. Likewise, it is important that infants responded unsystematically to novel label comprehension questions following baseline. This clarifies that infants' comprehension performance following labeling was not due to a bias to select the toy they had had in their possession, as that simple bias would have led to systematic selection of the target object in response to comprehension questions following the baseline phase as well.

   Still, infants' comprehension performance did differ somewhat across studies: in Study 1, only the older age group exhibited a tendency to learn the novel label, whereas in Study 2, infants at both ages learned the novel label. This difference might again be explained by procedural details: in Study 1, the experimenter ensured that infants were

paying attention to her before she labeled, whereas in Study 2, the experimenter ensured that infants were focused on the target object before she labeled. This might have made the label–object association easier to make in Study 2 than in Study 1. Further, infants in Study 2 were asked four novel label comprehension questions (as opposed to just two in Study 1), perhaps providing additional sensitivity to comprehension in Study 2.

Interestingly, Study 2's results indicate that infants sought more information about the speaker's object of reference in an ambiguous learning context, yet were also inclined to establish new label–object linkages in these same contexts. That is, Study 2's comprehension findings that infants linked the novel label with the target object to roughly the same degree regardless of the ambiguity of the learning context stand in contrast to infants' looking behavior, as infants looked more to the speaker in the ambiguous two-object context than the less ambiguous one-object context. In this sense, infants' comprehension performance was less indicative of their sensitivity to the referential ambiguity of the learning context than their looking behavior. The mechanism promoting infants to link the novel label with the target object when two objects were present is unclear; infants may have relied on something akin to an associative strategy—linking the label to the object closer to themselves—or they may have drawn the pragmatic inference that, given the lack of information to the contrary, the speaker was probably referring to the object closer to themselves. Although the precise mechanism cannot be known from the present data, it is striking that infants sought referential information in an ambiguous context but nevertheless established new label–object links in this same context.

In sum, the results of Study 2 bore out those of Study 1 with a larger sample and a modified procedure, and showed that already by 13 months, infants seek out information in the face of referential ambiguity. In what follows, we consider the implications of these findings.

General Discussion

One of the hallmarks of human 'cultural learning' (Tomasello, 1999), is arguably our ability to acquire knowledge about the world from conspecifics in a way that is efficient, effective, and appropriate (Csibra & Gergely, 2006). A critical element of this knowledge acquisition is the ability to recognize situations in which one requires more information. Prior work (Baldwin, 1991, 1993; Moses et al., 2001) has not shown unequivocally that infants possess this ability, and there has been significant reason to question whether they likely would (Baldwin & Moses, 1996). The current studies addressed this question by varying the level of ambiguity of the situation. Specifically, we examined how infants gather information about referents of novel words in situations where the referent is relatively obvious (just one possible referent) vs. ambiguous (two possible referents).

The results suggest that both 13- and 18-month-old infants recognize situations in which they require disambiguating information. This was apparent in Study 1, in which infants looked longer to the labeler, and in Study 2, in which infants directed more looks to the labeler in the presence of two rather than one novel object. Note that these differences were not apparent during the non-labeling baselines in either study. Our findings are thus consistent with the possibility that in referentially ambiguous situations, infants can seek out clarifying referential information. Thus, when infants hear a speaker produce a novel label, they do not simply orient to her voice and then follow

her gaze to the referent, but actively look to her because they need more information about the referent.

We argue that infants looked less to the speaker in the one object condition because they could infer that the novel label must refer to the only novel object present. Infants' comprehension performance in both studies provides support for this claim because it shows that indeed, infants (in the older age group in Study 1 but in both age groups in Study 2), did infer that the novel label referred to the novel object.

Our findings have implications for theories of word learning. Multiple interacting mechanisms have been proposed for young children's ability to infer meanings of new words (see Hollich, Hirsh-Pasek, & Golinkoff, 2000; Woodward & Markman, 1998). Some argue that children come to the word-learning situation with a set of assumptions that constrain the possible meanings children entertain (Markman & Wachtel, 1988; Waxman & Markow, 1995). Others argue that children use syntactic cues (Gelman & Markman, 1985; Naigles, 1990), or rely on general processes of attention and learning (Samuelson & Smith, 1998). Certain features of the input such as talk about here-and-now are also thought to make the meanings of new words less ambiguous (Clark & Wong, 2002). Our findings add to the literature by showing that children also rely on socio-pragmatic cues that indicate speakers' referential intentions (see also Akhtar, Carpenter, & Tomasello, 1996; Baldwin, 1993): when faced with ambiguity about the speaker's referential intent, infants actively seek clarifying gaze information before mapping a novel word onto a novel object.

The current studies focused on infants' behavior in word learning contexts, but an equally important context is social referencing, in which emotional (rather than linguistic) information is required. Some past work has explored why infants look to emoters in such contexts. Striano, Vaish, and Benigno (2006), for example, explored whether infants on a visual cliff look to their mother for information or for comfort. The authors varied the height of the visual cliff such that some infants faced a cliff of ambiguous height whereas others faced an unambiguously deep cliff. The reasoning was that if infants look for information, they should look more to their mother in the ambiguous case, whereas if they look for comfort, they should look more in the unambiguous case (see Baldwin & Moses, 1996). However, the results were equivocal, due perhaps to the special characteristics of the visual cliff, which the authors suggest might not be the right paradigm to examine this question (but see Stenberg & Hagekull, 2007, for similar results using a novel toy paradigm).

We created a simpler paradigm using ambiguous and unambiguous word learning situations, and found that infants do seek out information about novel words in ambiguous situations. This paradigm could be adopted to assess whether infants also seek out emotional information. Specifically, infants could be provided with emotional information about one vs. two novel objects. If infants look to the emoter more in the two objects case, this would indicate that infants robustly seek out information in multiple types of ambiguous situations, supporting the idea of a domain-general social information-gathering process (e.g., Diesendruck & Markson, 2001). If they do not look more in the two objects case, however, then perhaps infants are especially motivated to learn the referents of novel words but not yet of emotional information, or perhaps their ability to gather language-related information is somehow more advanced than their ability to gather other social information.

In conclusion, infants and young children learn entire systems of information with surprising ease, based upon a multitude of social and cognitive skills (Baldwin, 2000; Sabbagh & Baldwin, 2001; Tomasello, 1999). One critical skill involves recognizing

when they need more information. It has been proposed that through the first year and into the second year, infants are good information consumers but may not be good information seekers (Baldwin & Moses, 1996). However, in the present studies, we showed for the first time that 13- and 18-month-old infants possess some of the core, requisite skills to be good information seekers: They successfully identify at least some situations that require disambiguating information, and then actively seek out the required information in order to learn accurately. Although further skills, such as identifying knowledgeable vs. ignorant informers, likely still need to emerge in order for children to engage in true social information gathering (Baldwin & Moses, 1996), our findings nevertheless highlight the very early development of a critical social information-gathering skill. Such skill advances infants' knowledge acquisition by enabling them to more effectively capitalize on the vast knowledge that those around them possess.

References

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, *67*, 635–645.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*, 875–890.

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, *20*, 395–418.

Baldwin, D. A. (2000). Interpersonal understanding fuels knowledge acquisition. *Current Directions in Psychological Science*, *9*, 40–45.

Baldwin, D. A., & Moses, L. J. (1996). The ontogeny of social information gathering. *Child Development*, *67*, 1915–1939.

Clark, E. V., & Wong, A. D. W. (2002). Pragmatic directions about language use: Offers of words and relations. *Language in Society*, *31*, 181–212.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata, & M. H. Johnson (Eds.), *Processes of Change in Brain and Cognitive Development. Attention and Performance XXI* (pp. 249–274). Oxford: Oxford University Press.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, *37*, 630–641.

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics*, *21*, 95–116.

Gelman, S. A., & Markman, E. M. (1985). Implicit contrast in adjectives vs. nouns: Implications for word-learning in preschoolers. *Journal of Child Language*, *12*, 125–145.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*(3, Serial No. 262).

Mangold International GmbH. (2007). Interact (Version 8.0). [Computer software]. Arnstorf: Mangold Interactional GmbH.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology*, *20*, 121–157.

Moses, L. J., Baldwin, D. A., Rosicky, J. G., & Tidball, G. (2001). Evidence for referential understanding in the emotions domain at twelve and eighteen months. *Child Development*, *72*, 718–735.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*, 357–374.

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*, 1054–1070.

Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, & Tomasello. *Child Development*, *69*, 94–104.

Stenberg, G., & Hagekull, B. (2007). Infant looking behavior in ambiguous situations: Social referencing or attachment behavior? *Infancy*, *11*, 111–129.

Striano, T., Vaish, A., & Benigno, J. P. (2006). The meaning of infants' looks: Information-seeking and comfort-seeking? *British Journal of Developmental Psychology*, *24*, 615–630.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*, 257–302.

Woodward, A. L., & Markman, E. M. (1998). Early word learning. In W. Damon (Series Ed.), D. Kuhn, & R. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception and language* (5th ed., pp. 371–420). New York: Wiley.

Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, *30*, 553–566.